Title

Cloud base height estimates from sky imagery and a network of pyranometers

Authors

Wang, GC
Urquhart, B
Kleissl, J

Publication Date

2019-05-15

DOI

10.1016/j.solener.2019.03.101

Peer reviewed

Cloud base height estimates from sky imagery and a network of pyranometers

Guang Wang, Bryan Urquhart, Jan Kleissl
Center for Renewable Resources and Integration, Department of Mechanical and Aerospace Engineering, University of California, San Diego, United States

Abstract

Cloud base height (CBH) is an important parameter for physics-based high resolution solar radiation modeling. In sky imager-based forecasts, a ceilometer or stereographic setup is needed to derive the CBH; otherwise erroneous CBHs lead to incorrect physical cloud velocity and incorrect projection of cloud shadows, causing solar power forecast errors due to incorrect shadow positions and timing of shadowing events. In this paper, two methods to estimate cloud base height from a single sky imager and distributed ground solar irradiance measurements are proposed. The first method (Time Series Correlation, denoted as "TSC") is based upon the correlation between ground-observed global horizontal irradiance (GHI) time series and a modeled GHI time series generated from a sequence of sky images geo-rectified to a candidate set of CBH. The estimated CBH is taken as the candidate that produces the highest correlation coefficient. The second method (Geometric Cloud Shadow Edge, denoted as "GCSE") integrates a numerical ramp detection method for ground-observed GHI time series with solar and cloud geometry applied to cloud edges in a sky image. CBH are benchmarked against a collocated ceilometer and stereographically estimated CBH from two sky imagers for 15 minute median-filtered CBHs. Over 30 days covering all seasons, the TSC method performs similarly to the GCSE method with nRMSD of 18.9% versus 20.8%. A key limitation of both proposed methods is the requirement of sufficient variation in GHI to enable reliable correlation and ramp detection. The advantage of the two proposed methods is that they can be applied when measurements from only a single sky imager and pyranometers are available.

Keywords: Cloud base height; Sky imager; Irradiance ramp detection; Short-term solar forecasting

33

34 # Nomenclature

$\mathrm{GHI}_i(t;H)$	GHI simulated using USI imagery at a given CBH	t_o	Current time
$\mathrm{GHI}_i^{\mathrm{obs}}(t)$	GHI from the pyranometer at station i	$\hat{\mathbf{u}}$	Cloud pixel speed [pixel s^{-1}]
$\mathrm{GHI}_i^{\mathrm{csk}}(t)$	GHI from clear sky model	$\mathbf{u}$	Cloud speed [m s^{-1}]
H	Cloud base height (CBH)	$\mathbf{x}_c$	Intersection of cloud motion line and cloud boundary [m]
H^{ceil}	CBH measured by ceilometer	$\hat{\mathbf{x}}_c$	Intersection of cloud motion line and cloud boundary [pixels]
H_j	CBH candidate	$\mathbf{x}_g$	Vector describing ground station location [m]
H^{model}	CBH estimate from Time Series Correlation and Geometric Cloud Shadow Edge methods	$\hat{\mathbf{x}}_g$	Vector describing ground station location [pixels]
h	Sky imager elevation	$\mathbf{x}_s$	Intersection of solar beam and cloud map
K	Number of samples in 20 minutes at 30 second intervals	ΔH	Cloud base height error [m]
kt	Clear sky index	Δt	Cloud travel time, a time difference between given initial timestamp t_i and start of next down ramp
M	Number of modeled CBH values	Δt_f	Forecast time step
MBE	Mean bias error	$\Delta\mathbf{x}$	Cloud shadow horizontal shift corresponding to cloud base height vertical shift ΔH [m]
n_p	Number of cloud map pixels in one dimension	$\Delta\hat{\mathbf{x}}$	Cloud displacement in the sky image [pixels]
N	Total number of available ground sites	$\Delta\mathbf{x}_c$	Cloud projection error
nMBE	Normalized mean bias error	θ	Zenith coordinates of a pixel in the sky image
nRMSD	Normalized root mean square difference	θ_m	Sky imager field of view in degrees from the vertical
O	Sky imager position	θ_s	Solar zenith angle
R	Length of cloud map in one dimension	λ	Distance along a ray from observation point
R_{ij}	Correlation coefficient between $\mathrm{GHI}_i(t;H_j)$ and $\mathrm{GHI}_i^{\mathrm{obs}}(t)$ at site i for CBH H_j	μ	Cloud velocity scaling factor
R_j	Correlation coefficient averaged over all sites at CBH H_j	μ_i^{obs}	Mean of $\mathrm{GHI}_i^{\mathrm{obs}}(t)$
RMSD	Root mean square difference	μ_{ij}	Mean of $\mathrm{GHI}_i(t;H_j)$
t	Time	ϕ	Azimuth coordinates of a pixel in the sky image
t_i	Initial timestamp used to compute Δt	ϕ_s	Solar azimuth angle

35

1. Introduction

1.1 Impact of CBH on Intra-hour Solar Power Forecasting with a Sky Imager

CBH plays a vital role in intra-hour solar power forecasting. For typical mid-latitude solar zenith angles of 45°, a difference of 100 m in CBH causes a 100 m translation of the cloud shadow on the ground (Eq. (1). In addition, since opaque clouds typically have a clear sky index of 0.4, local power output forecast errors of 60% of clear sky production levels (Martinez-Anido et al., 2016) are common if a CBH error causes the wrong sky condition (clear or cloudy) to be forecast. Thus, accurate CBH estimation is critical for predicting local power ramps over short time scales.

For sky imager solar forecasts that are based on the geometry between the sun, clouds, and ground, CBH is required for mapping the cloud field from sky images to the atmosphere and then projecting to the ground. Specifically, the mapping process consists of three geometry steps: 1) projection of the clouds in the sky image into a plane in the sky (termed "cloud map", see Section 2.3) at the CBH; 2) forward motion of the cloud map in time; 3) projection of cloud map onto the ground. Thus, an erroneous CBH leads to three different scaling errors listed below (see the nomenclature for variable definitions and Section 3.2. for derivations:

1) The cloud projection error is:

$$\Delta\mathbf{x}_c = \Delta H \cdot (tan\,\theta\, sin\,\phi\,, tan\,\theta\, cos\,\phi\,, 1)^\top \quad (1)$$

where $\Delta\mathbf{x}_c$ is a 3D-vector describing position error for a given CBH error ΔH, and (θ, ϕ) are respectively the zenith and azimuth pointing angles corresponding to a pixel obtained using pixel coordinates (refer to Figure 4 later) and the camera geometric calibration (e.g. Urquhart et al. 2016). ΔH linearly scales cloud horizontal position in the radial direction and stretches or shrinks the cloud about a center point at the sky imager, and the scaling error is more sensitive to ΔH at farther spatial distance (outer pixels) caused by the nonlinear effect of $tan\,\theta$.

2) Physical cloud velocity error. Because the cloud velocity derived from sky image is in units of pixels, a conversion to actual cloud velocity in units of m/s requires scaling the pixel velocity with CBH, resulting in a linear scaling error by ΔH.

3) Cloud shadow projection error. When the cloud map is advected and projected onto the ground, the vertical shift ΔH causes a uniform horizontal shift $|\Delta\mathbf{x}|$ in shadow position following the expression:

$$|\Delta \mathbf{x}| = \Delta H \, tan\, \theta_s, \quad (2)$$

which is exaggerated at larger solar zenith angles θ_s. Thus, CBH errors also cause shadows or sunlight to be predicted at locations that are shifted further as the distance from the sky imager increases.

1.2 **CBH Measurement Techniques**

CBH can be measured directly using in-situ and remote sensing instruments such as radiosondes (Wang & Rossow, 1995), ceilometers (Gaumet et al., 1998; Martucci et al., 2010), and satellites (Hutchison et al., 2006). A radiosonde is a battery-powered telemetry instrument package that vertically profiles the atmosphere as the balloon ascends, yielding CBH estimates. Although the CBH measurements from a radiosonde are accurate, the observations are usually taken at most twice daily and at discrete and sparse locations, making them unsuitable for use in intra-hour solar energy forecasting. Ceilometers are the most common CBH observational tool and are regularly installed at airports and meteorological aerodrome reports (METAR) stations. It emits a pulsed near-infrared vertical laser beam and measures a vertical profile of atmospheric backscatter from which CBH is derived. Since ceilometers can be expensive, they have limited application outside of airports in most countries except in the UK, where ceilometer is a standard component of weather stations.

Indirect CBH measurements using ground based thermal infrared cameras (Shaw and Nugent, 2013; Liu et al., 2015) and derived data from remote-sensing techniques such as spectroradiometers (Hutchison et al., 2006) are also feasible. The assumption that clouds are blackbodies usually leads to an overestimation of CBH derived by infrared cloud imagers (Liu et al., 2015). Satellite-measured cloud top near-infrared radiance (Dessler et al., 2006) or measured cloud top temperature with an atmospheric temperature profile (Prata & Turner, 1997) can be used to obtain cloud top height with wide spatial coverage, but CBH is difficult to detect from satellites and time delays in data dissemination limit its application in short-term solar power forecasting. Numerical weather prediction offers another alternative to obtain CBH (Killius et al., 2015).

CBH can also be obtained from sky imagery. The application of stereogrammetric techniques using two sky imagers was investigated by Allmen and Kegelmeyer (1996) and Kassianov et al. (2005). Nguyen and Kleissl (2014, referred to as NK14) further generalized and improved accuracy and computational efficiency of the approach introduced by Kassianov et al. (2005) for (binocular) stereographic CBH estimation: a two-dimensional (2D) georeferenced projection is

used to overlay images from each camera. The CBH is the cloud height associated with the minimum normalized matching error, which implicitly assumes a single cloud layer. More sophisticated stereo-vision techniques can offer cloud base height estimate in 3D coordinates using the standard technique of matching image patches along epipolar curves (Allmen and Kegelmyer, 1996; NK14; Kleissl et al., 2016). These methods are computationally intensive and provide high spatial resolution CBH within a pair of images. The stereographic method requires at least two sky imagers and accurate geometric calibration of the imaging system (e.g. Urquhart et al., 2016). Wang et al. (2016) and Kuhn et al. (2018a; 2018b) demonstrated that CBH can be obtained from a single sky imager and an independent measurement of cloud speed. Because angular cloud speed determined from sky images is proportional to cloud speed and CBH, CBH can be derived from a collocated cloud speed sensor (Fung et al., 2014) and sky imager. In Wang et al. (2016) and for the same location as in this paper, typical daily root mean square differences were 126 m or 17% of the observed CBH. But the raw (instantaneous) CBH measurements need to be filtered to derive a robust CBH, which makes CBH outputs infrequent (one CBH output every 50 sec for 27 partly cloudy days and every 250 sec for 21 overcast days, on average).

1.3 **Objectives and Structure of the Paper**

CBH is a required input for some sky imager-based short-term solar power forecasting variants (Chow et al., 2011; Schmidt et al., 2016). The variety of methods presented in Section 1.2 can produce accurate CBH information at different temporal and spatial scales, however either equipment or operating costs are prohibitive, or computational requirements are high, or the temporal resolution is insufficient for intra-hour solar power forecasting.

Cameras are ubiquitous and low cost, and nearly every solar power installation has pyranometers and PV energy meters. Therefore, existing and low cost infrastructure provides an opportunity to estimate cloud height as an ancillary product if the irradiance distribution on the ground is measured in space and time. Thus, the objective of this work is to provide a low-cost alternative to estimate CBH using such irradiance measurements and a single sky-pointing camera. CBH is estimated using two related methods requiring a single sky imager and irradiance sensors distributed within the footprint of the sky imager, i.e. within the camera's field of view. Both methods are new and have not been been presented before. In the first method, CBH is estimated by correlating ground-observed GHI measured using a set of pyranometers with GHI modeled using a sky imager irradiance forecast (Chow et al., 2011). Modeled GHI time series are generated from a sequence of sky images geo-rectified to a candidate set of CBH. The second method estimates CBH by matching ramp event timings from pyranometer-measured GHI to

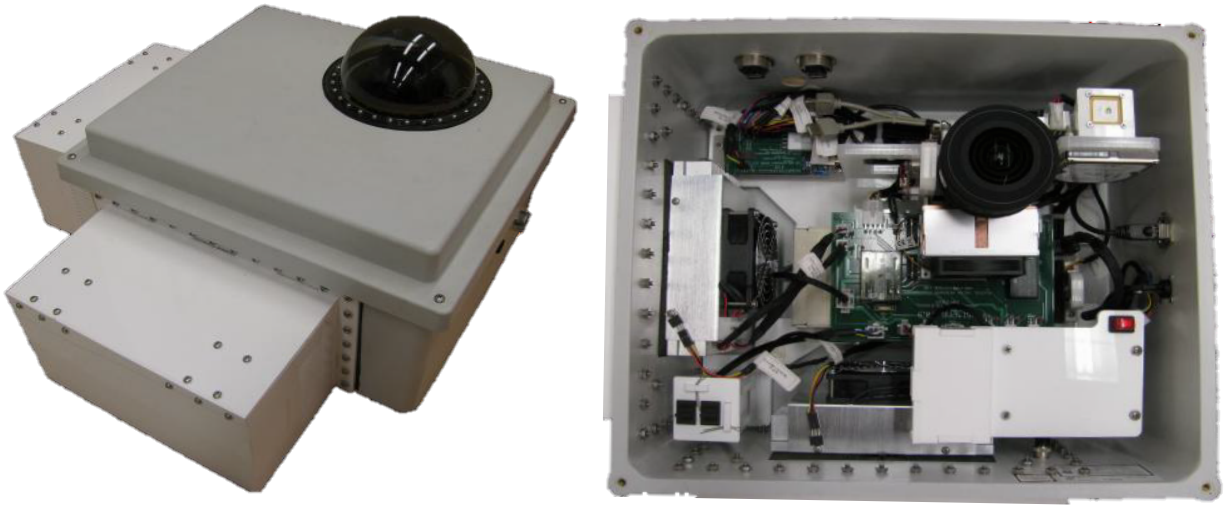
cloud shadow arrival times derived from cloud geometry and sun triangularization adapted to sky imagery. The presentation of the latter method provides a new mathematical description of the forecast approach used in Chow et al. (2011).

This paper is organized as follows. The measurement equipment, including the sky imaging system and forecasting procedure, is briefly described in Section 2. Section 3 introduces the CBH estimation methods. Section 4 presents the overall performance in a set of 30 days, and then validates CBH from both methods against ceilometer data and the NK14 stereographic method in a case study. Section 5 provides detailed discussion regarding the performance and limitation of the proposed methods. Finally Section 6 provides conclusions and future work.

2. Experimental Data and Sky Imager Forecast Procedure

2.1 **Ground Measurements**

The University of California, San Diego (UCSD) designed and developed a sky imager system specifically for short-term solar power forecasting applications (Fig. 1, Urquhart et al., 2013). The UCSD Sky Imager (USI) features a high-quality image sensor and lens contained in a thermally controlled, compact environmental housing, and capture software employing a high dynamic range (HDR) imaging technique. The USI uses an Allied Vision GE-2040C camera which has a 15.15 × 15.15 mm ON Semiconductor KAI-04022 CCD sensor (originally developed by Kodak). The Sigma 4.5 mm focal length fisheye lens provides a 180 degree field of view with 1748 × 1748 pixels covering the sky hemisphere. Thermal stability of the camera is achieved using two thermoelectric coolers for the entire enclosure, a copper heat sink, and a fan attached to the camera to keep it at the ambient enclosure temperature. The dome on the USI is a 1.6 mm thick, neutral density (ND2) acrylic hemisphere with a UV protective coating. Additional information can be found in Urquhart et al. (2015). The USI used in this analysis is installed next to one of the six pyranometers shown in Figure 2 and Table 1.



(a) (b)

Figure 1: The University of California, San Diego Sky Imager (USI). (a) Outer view showing the enclosure with dome and white radiation shields for the coolers; (b) a top view of the open system showing the components inside the enclosure.

158

159 GHI data sampled at 1 Hz is obtained from six weather stations with Li-COR 200SZ
160 pyranometers installed at the locations shown in Figure 2 and Table 1. In addition, a Vaisala
161 CT25K ceilometer located on EBU2 computes CBH every 20 seconds from backscatter returns.
162 Due to the small sampling area (a small <0.1° cone above the ceilometer), the heterogeneity of
163 cloud field, as well as cloud formation and movement, the 20-second ceilometer output is not
164 always representative of the CBH in the field of view of the sky imager. Therefore, consistent with
165 NK14, a 15-minute median filter is applied to ceilometer measurements prior to comparison with
166 the proposed methods.

167

168

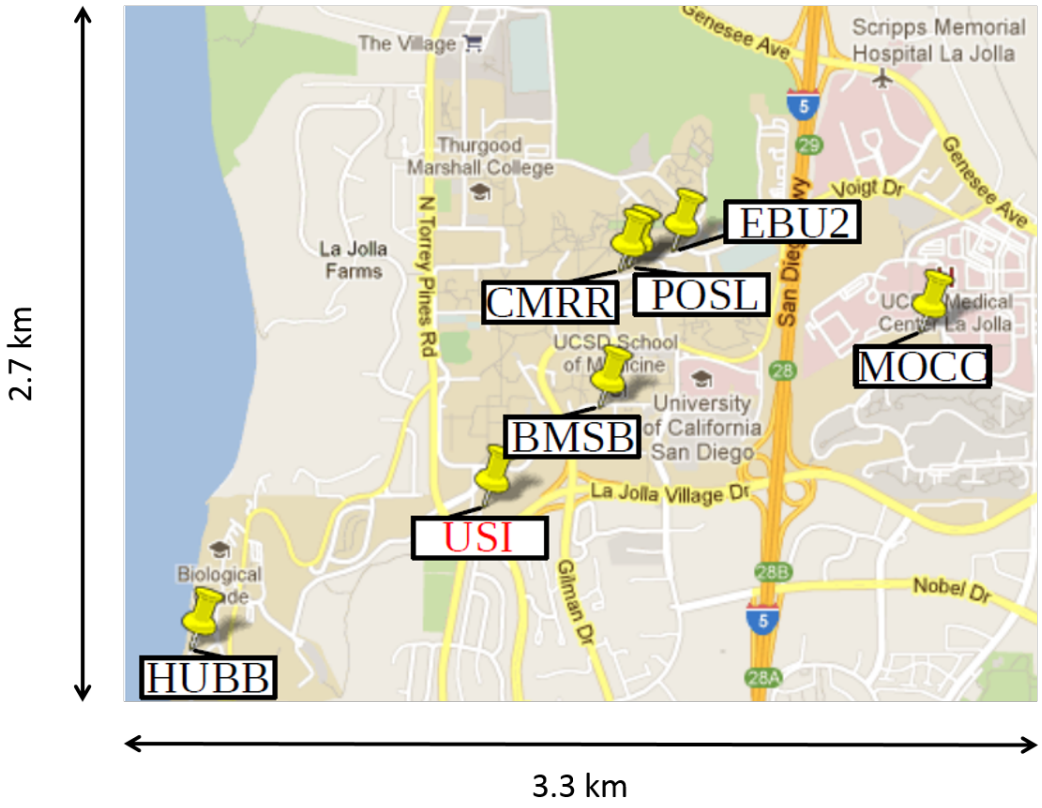


Figure 2: Locations of the six pyranometers and the USI on the UCSD campus. The ceilometer is located on EBU2. Reprinted with permission from Yang et al. 2014. © Google Maps.

Table 1: Locations of USI and pyranometers used for CBH estimation and their respective distances to the USI. (*re-tabulated with permission from Yang et al. 2014)*

Station Name	Latitude	Longitude	Altitude (m MSL)	Distance to USI (m)
USI	32.8722	-117.2410	140	-
BMSB	32.8758	-117.2362	111	603
CMRR	32.8806	-117.2353	111	1074
EBU2	32.8813	-117.2330	101	1257
HUBB	32.8672	-117.2534	24	1288
MOCC	32.8784	-117.2225	103	1857
POSL	32.8807	-117.2350	110	1103

2.2 **Evaluation Dataset**

The CBH estimation methods are evaluated using two different sets of CBH measurements: (1) an on-site ceilometer on 33 days and (2) the NK14 2D stereography method on 3 days. Thirty-

three cloudy days from 2012 to 2016 were selected based on the following criteria:

1) Data availability from sky imager, ceilometer and pyranometers.
2) Cloudy conditions: clear and rainy days were excluded.
3) Cloud type: opaque clouds such as stratocumulus, cumulus, and stratus, since they are most relevant to solar forecasting of GHI that is the subject of this paper.
4) Cloud height predominantly less than 1000 m. Four days were chosen with cloud heights greater than 1000 m
5) Lack of rain: less than 2 hours of rain

Finally, only time periods with solar zenith angles less than 75° are considered. Moreover, during an intensive operating period in 2012, two sky imagers were installed, which allowed 2D stereography to be applied to four days, as reported in NK14. December 14, 2012 was characterized by broken stratocumulus clouds above a few cumulus clouds. On December 26, a single layer of low scattered cumulus clouds was observed. December 29 was overcast with stratus clouds. Jan 1, 2013 analyzed in NK14, was not included in this paper because several station outages limited GHI measurements to only two stations.

2.3 **Sky Imager Forecast Procedure**

The USI can be used to geolocate clouds, to measure cloud angular velocity, and to track cloud motion (Chow et al., 2011; Chow et al., 2015). These measurements are then used to forecast future cloud locations up to 15 minutes ahead. The forecast procedure is outlined in the flow chart of Figure 3. A brief overview of the USI forecast procedure is given in the remainder of this section. For more information, the reader is referred to Chow et al. (2011), Ghonima et al. (2012), Urquhart et al. (2013), and Yang et al. (2014). Similar sky imager systems and forecast procedures can be found in Cazorla et al. (2010); Marquez and Coimbra (2013), and Schmidt et al. (2016).

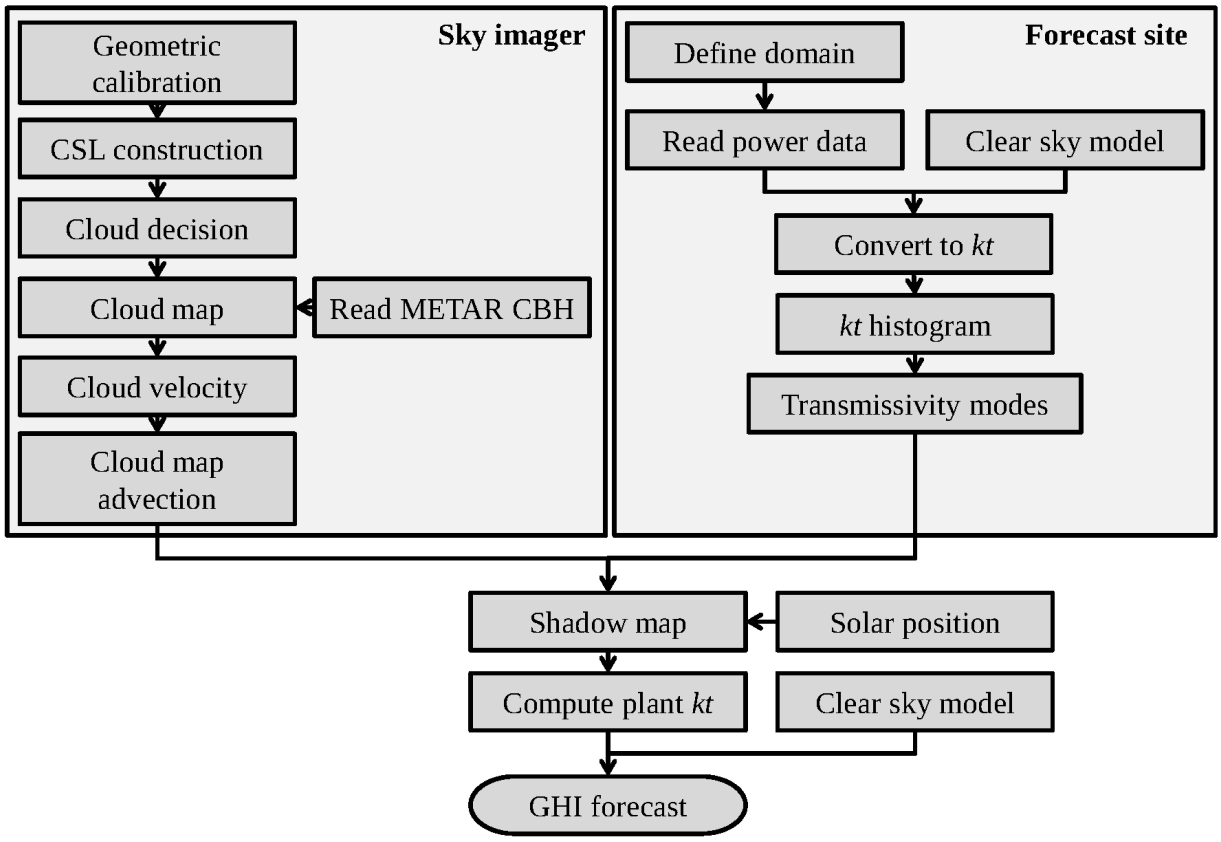


Figure 3: Flowchart of USI forecast procedure. Sky image processing (left) is combined with the clear sky index (*kt*) from local ground observations (right) to produce spatial irradiance forecasts. *(reprinted with permission from Yang et al. (2014))*

Based on images taken every 30 seconds, cloudy pixels are detected and using lens-camera geometry, images are transformed to a rectified planar grid (Allmen and Kegelmeyer, 1997). CBH is then used to register each pixel to a latitude, longitude, and altitude (geo-rectification, Chow et al., 2011). The resulting geo-referenced map of clouds is termed the "cloud map", which is a planar mapping of cloud position at a specified altitude above the forecast site. The cloud map at the current time $t = t_o$ yields the real time solar irradiance *forecast* (which would be sensibly called "nowcast" even though commonly the word "nowcast" is associated with minutes-ahead forecast), while future cloud positions $(t > t_o)$ are determined through cloud advection at discrete time steps delivering the short-term solar irradiance forecast. The ability to resolve the horizontal cloud structure near the horizon is limited due to perspective effects (look vectors are nearly parallel to horizontal cloud base) and due to the longer distance to the clouds, causing a single pixel to subtend a much larger projected area. Both of these factors introduce errors when using the perimeter of the cloud map (more discussion in Section 5.3).

Cloud pixel velocity is obtained by applying a cross-correlation method to the red-blue ratio of two consecutive sky images. The cloud speed $\mathbf{u}$ [m s^{-1}] is then calculated from cloud pixel velocity $\hat{\mathbf{u}}$ [pixel s^{-1}] using a scaling factor μ, which is a function of CBH as:

$$\mathbf{u} = \frac{1}{\mu(H)}\hat{\mathbf{u}} = \frac{1}{\mu}\frac{\Delta\hat{\mathbf{x}}}{\Delta t_f}, \tag{3}$$

where $\Delta\hat{\mathbf{x}}$ is the cloud displacement in the image, Δt_f is the image capture interval (here also equal to the forecast time step), and the ^ indicates units of pixels. Equation (9**Error! Reference source not found.** in Section 3.2 gives the expression for $\mu(H)$. The cloud velocity is then used to advect the planar cloud map to generate cloud position forecasts for each forecast horizon. Since the distance from the sun to the Earth is much larger than the distance from the clouds to the Earth (i.e. the direct solar beam for locations on Earth is essentially parallel), cloud shadow speed is essentially identical to cloud speed.

The forecast procedure used in this work is developed for a single sky imager. The default CBH source for a single sky imager is METAR. METAR stations, which use a ceilometer, report high quality CBH data but are limited in temporal resolution (typically hourly reports) and are spatially sparse. Therefore, spatial variability in cloud cover causes differences between CBH at the sky imager location and the nearest METAR station. These limitations are the main movtivation for this work.

3. Methods for CBH Estimation

A Time Series Correlation (TSC) method and a Geometric Cloud Shadow Edge (GCSE) method will be introduced in this section. Both methods only require a single sky imager and time-synchronized measurements of GHI or solar power output at surrounding stations. For TSC, at each ground station GHI is simulated for a set of CBHs and cross-correlated with GHI measurements at the corresponding ground sites. For GCSE, cloud arrival and departure times are determined from the GHI time series using ramp detection. CBH is then derived by matching these detected cloud arrival times with cloud arrival times simulated using USI cloud imagery and cloud position forecasts.

3.1 **TSC Method**

Most of the large-magnitude variability in GHI time series is introduced by cloud shadows

approaching or departing a location. In fact, as described in Wang et al. (2016), cloud shading events implicitly contain CBH information: the duration of the shading event is proportional to the length of cloud (and cloud shadow) in the direction of cloud motion (cloud velocity assumed to be constant). Using an independent cloud speed measurement (e.g. Bosch and Kleissl, 2013; Bosch et al., 2013) along with cloud pixel speed estimated in the USI forecast procedure (Section 2.3), CBH can be derived based on Eq. **Error! Reference source not found.**.

TSC estimates CBH using a grid search performed over a set of candidate CBH values H_j. For each ground measurement station (indexed by $i = 1 \dots N$), GHI is modeled over the last 20 minutes for each H_j ($j = 1 \dots M$) using USI nowcasts from a 20 min sequence of geo-rectified sky images captured at sampling rate of 30 sec (i.e. a total of $K = 41$ image samples). For each station, the correlation coefficient R_{ij} is computed between each modeled GHI time series $\mathrm{GHI}_i(t; H_j)$ and the observed GHI time series $\mathrm{GHI}_i^{\mathrm{obs}}(t)$:

$$R_{ij} = \frac{1}{K\sigma_i^{\mathrm{obs}}\sigma_{ij}} \sum_{k=1}^{K} \left[\mathrm{GHI}_i^{\mathrm{obs}}(t_o + k\Delta t_f) - \mu_i^{\mathrm{obs}}\right]\left[\mathrm{GHI}_i(t_o + k\Delta t_f; H_j) - \mu_{ij}\right], \quad (4)$$

where μ_i^{obs} and μ_{ij} are the means of $\mathrm{GHI}_i^{\mathrm{obs}}(t)$ and $\mathrm{GHI}_i(t; H_j)$ over the K samples, respectively, and σ_i^{obs} and σ_{ij} are the corresponding standard deviations. For each of the N stations, this yields M correlation coefficients. The coefficients are then averaged across stations for each value of H_j to generate a correlation score for each CBH candidate:

$$R_j = \frac{1}{N}\sum_{i=1}^{N} R_{ij}\,. \quad (5)$$

Initially, a weighting scheme using the inverse sky imager to weather station distance was applied, however performance was similar, and thus only a simple average is used here. After R_j has been computed for all CBH candidates H_j, the CBH candidate corresponding to the largest correlation score R_j is selected as the CBH estimate.

Theoretically, TSC can yield a CBH every 30 seconds (i.e. sampling rate of sky images) because a correlation can always be established. However, since CBH in clear or rainy conditions is irrelevant to solar forecasting, TSC results with correlation coefficients below 0.5 are excluded. Moreover, the performance of TSC degrades in homogenous cloud cover or clear conditions because the variations in the time series are small and correlation between modeled and

measured GHI is expected to be similar for all CBH candidates. As further discussed in Section 5, under these conditions, using the maximum correlation is not a reliable way to estimate CBH. Fortunately, for solar power forecasting applications, in cases of uniform sky cover, the impact of CBH error is mitigated.

3.2 **GCSE Method**

3.2.1 **Cloud Shadow Geometry**

The coordinate system origin is the sky imager position. The coordinate axes are aligned such that x is positive east, y is positive north, and z is positive up, and earth curvature effects are ignored. The location of a ground station in this coordinate system is then $\mathbf{x}_g = \left(x_g, y_g, z_g\right)^\top$, where $\top$ indicates transpose. The ray pointing to the sun from point $\mathbf{x}_g$ can be parameterized as:

$$\mathbf{x}_s(\lambda) = \mathbf{x}_g + \lambda \begin{bmatrix} \sin\theta_s \sin\phi_s \\ \sin\theta_s \cos\phi_s \\ \cos\theta_s \end{bmatrix}, \tag{6}$$

where θ_s is the solar zenith angle, ϕ_s is the solar azimuth angle, and λ is the distance from $\mathbf{x}_g$ towards $\mathbf{x}_s$ in meters. Assuming a planar layer of clouds, we can compute the intersection of $\mathbf{x}_s(\lambda)$ with the clouds by setting the z-coordinate to the CBH above the sky imager: $x_{s,z}(\lambda) = x_{g,z} + \lambda\cos\theta_s = H - h$, where H is the cloud base height and h is the height of the sky imager (both heights referenced above ground level [AGL]). This gives $\lambda = \left(H - h - x_{g,z}\right)\sec\theta_s$; the point $\mathbf{x}_s$ in the cloud layer is then:

$$\mathbf{x}_s = \mathbf{x}_g + \left(H - h - x_{g,z}\right)\begin{bmatrix} \tan\theta_s \sin\phi_s \\ \tan\theta_s \cos\phi_s \\ 1 \end{bmatrix} = \mathbf{x}_g + \left(H - h - x_{g,z}\right)\mathbf{s}\,, \tag{7}$$

where $\mathbf{s} = (\tan\theta_s \sin\phi_s\,, \tan\theta_s \cos\phi_s\,, 1)^\top$. Figure 4a illustrates this geometric relation when sky imager position O, $\mathbf{x}_g$, and $\mathbf{x}_s$ are coplanar in azimuth (although in general they are not coplanar) and Fig. 4b shows a top-down-view of the geometric configuration.

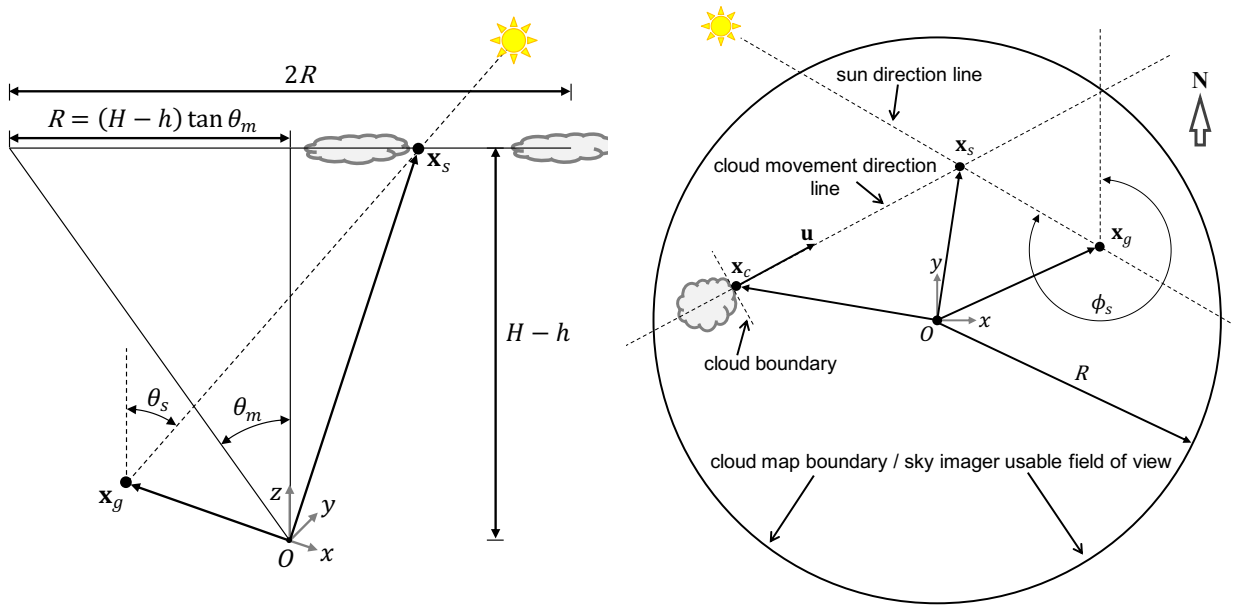


Figure 4: (a) Cross-section and (b) plan view of the geometric relationship between sky imager position O, a ground station $\mathbf{x}_g$ and the cloud intersection point $\mathbf{x}_s$. (Note: to improve the illustration clarity, $\mathbf{x}_g$ and $\mathbf{x}_s$ are shown in different relative locations in each subfigure)

Depending on the spatial configuration of the cloud field, at any given time clouds may or may not be present at point $\mathbf{x}_s$, which is the point at which clouds must be present to shade the station located at $\mathbf{x}_g$ (To shade a sensor at $\mathbf{x}_g$, clouds may actually be anywhere along $\mathbf{x}_s(\lambda)$, but again we are assuming a planar cloud field at height H). Assuming a constant cloud velocity $\mathbf{u} = (u_x, u_y, 0)^\top$, we estimate the current position of a cloud $\mathbf{x}_c$ that will move to point $\mathbf{x}_s$ in Δt seconds:

$$\mathbf{x}_c(H, \Delta t) = \mathbf{x}_s(H) - \Delta t \mathbf{u}\,, \tag{8}$$

where the meaning of the input argument to $\mathbf{x}_s$ has been changed from slant distance λ to CBH H following Eq. **Error! Reference source not found.**. Hereinafter, Δt is referred to as the cloud travel time (Eq. 11**Error! Reference source not found.**).

To search the image for clouds that could potentially cause shadowing of the sensor, we search the surface $\mathbf{x}_c(H, \Delta t)$ parameterized by H and Δt. This requires the following conversion from space coordinates to image coordinates. The usable field of view of the sky imager for cloud imaging is $2\theta_m$, and the corresponding width of the cloud map is $2R = 2(H - h)\tan\theta_m$. The number of pixels spanning the cloud map diameter is set to n_p (The cloud map is an 'undistorted'

311 plane-projected version of the original distorted image, taking into account the camera calibration).
312 The projection requires interpolation of the image and n_p can be set to a suitable value based on
313 the footprint of the sky image. In this paper, n_p = 1251 is the default value used in our sky imager
314 forecast algorithm. The conversion from units of meters to pixels is then:
315

$$\mu(H) = \frac{n_p}{2(H-h)\tan\theta_m}\left[\frac{\text{pixels}}{\text{meter}}\right]. \tag{9}$$

316
317 Combining Eqs. **Error! Reference source not found.** and **Error! Reference source not found.**
318 and multiplying by μ gives:
319

$$\hat{\mathbf{x}}_c(H,\Delta t) = \hat{\mathbf{x}}_g + \frac{n_p}{2}\frac{(H-h-x_{g,z})}{(H-h)\tan\theta_m}\mathbf{s} - \Delta t\hat{\mathbf{u}}\,, \tag{10}$$

320
321 where the ^ indicates coordinates have been converted to units of pixels. When $|\hat{\mathbf{x}}_c(H,\Delta t)| >$
322 $n_p/2$, the cloud point is outside of the cloud map and the cloud state cannot be retrieved (i.e. it is
323 outside of the sky imager's usable field of view). Additionally, we only consider cases where $|\hat{\mathbf{x}}_g| \leq$
324 $n_p/2$, as $|\hat{\mathbf{x}}_g| > n_p/2$ occurs if the station is outside the cloud map because H is too low, shrinking
325 the cloud map (i.e. R is small). When the latter criterion is not met, it is possible that the shadow
326 projection of the cloud map may still encompass the station, however for $|\hat{\mathbf{x}}_g| > n_p/2$ the station
327 is "far" and the reliability of the results is questionable. Interestingly, setting $\hat{\mathbf{x}}_g$ to the sky imager
328 location $(0,0,0)^\top$ shows that forecasts at the sky imager location do not depend on CBH. Using
329 Eq. (10, we can solve for the cloud travel time:

$$\Delta t = \frac{1}{|\hat{\mathbf{u}}|}\left|\hat{\mathbf{x}}_g - \hat{\mathbf{x}}_c(H,t) + \frac{n_p}{2}\frac{(H-h-x_{g,z})}{(H-h)\tan\theta_m}\mathbf{s}\right|. \tag{11}$$

330

331 ### 3.2.2 **Ramp Detection**

332 A ramp detection procedure is used to determine the start of down ramps in the ground-
333 observed GHI data. Down ramps are associated with cloud edge arrival times, and thus locating
334 down ramps by an edge detection method allows timing the expected passage of a cloud edge
335 over the station. Note that there are many edge detection methods available for 1D data such as
336 canny edge detection and sobel operator. Because GCSE depends on finding significant cloud-
337 edge induced fluctuations in the irradiance timeseries, we believe that large ramp events could

be found by any edge detection method. Thus, to remove the dependence on external algorithms, we choose to develop our own edge detection method.

Figure 5 presents a case study where our own detection process described below is applied to a GHI time series. Precise ramp timings require a high sampling rate, and in this analysis a 1 Hz dataset is used. Ground-observed GHI, sampled at 1 Hz at each station, is converted to clear sky index using the Kasten clear sky model (improved and described by Ineichen and Perez, 2002). A Gaussian filter is then applied to smooth the data (top subplot). The size of the filtering window is selected as 10 min, an empirical tradeoff value between effective noise reduction and signal shape preservation. Consistent with convention, the filter width is set to 3 standard deviation which comes out to 100 seconds. At each time step in the smoothed series, we compute the maximum difference in clear sky index between the current data point and any subsequent point within 90 seconds yielding a time series of maximum ramp magnitudes (blue and green curve in the bottom subplot). All such ramp points with a clear sky index change in magnitude of greater than 0.3 (30% clear sky index ramp) are collected (red). Then local extrema are located with the MATLAB implementation of Findpeaks[1], which usually gives a single time instant corresponding to the start time of each large ramp (black lines). When more than one ramp extremum is found per ramp, the point with greater associated ramp magnitude is selected. Finally, because sometime large ramps exhibit non-monotonic characteristics, causing the detected start time to deviate, the ramp event start time is corrected if there is a local maximum in the clear sky index within 5 seconds from the detected time instant (refer to Section 5.2.1 for more details).

[1] https://www.mathworks.com/help/signal/ref/findpeaks.html, accessed March, 2019.

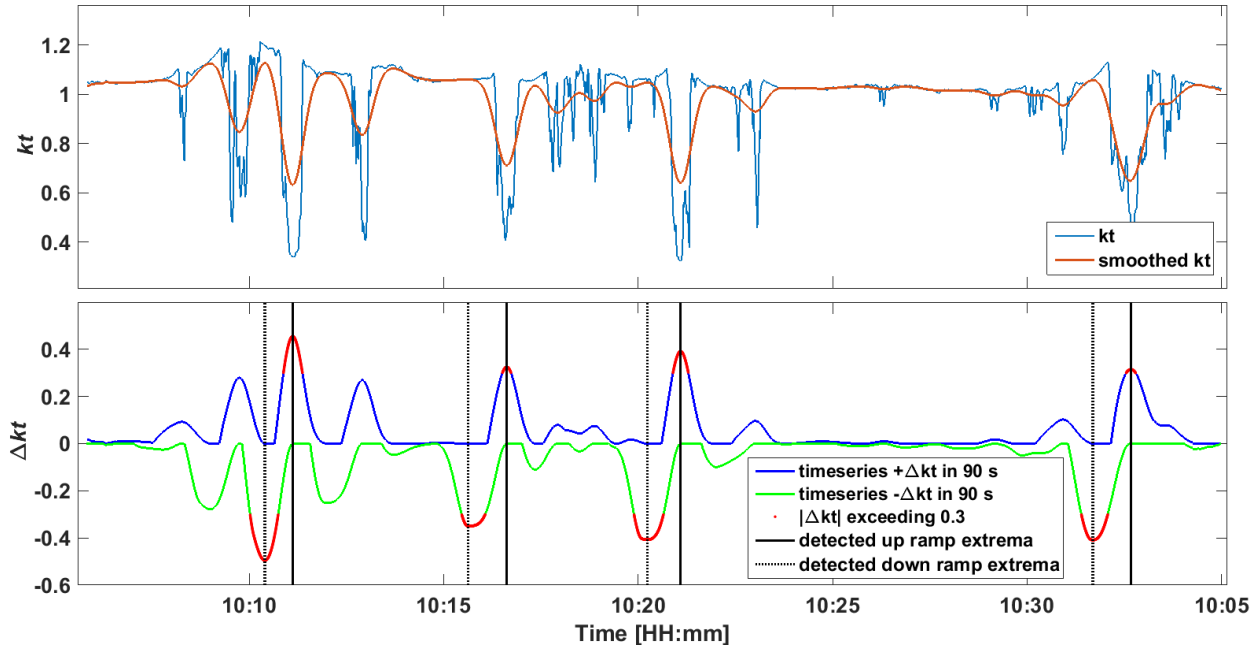


Figure 5: Illustration of the procedure to detect ramps in the normalized time series GHI (*kt*). Top: The time series *kt* is smoothed by a Gaussian filter with filter width of 10 min and standard deviation of 100 sec. Bottom: The maximum difference in *kt* between within time window of 90 sec is computed, resulting in time series ramp points of Δkt (blue and green). The points with an associated ramp magnitude of less than 0.3 are excluded and the remaining points are kept (red). The local extrema are located by MATLAB implementation of Findpeaks (black and dashed black line).

Figure 6 illustrates the outcome of a real execution of the procedure in Figure 5 to both BMSB and EBU2 stations (refer to Figure 2 for station name and location). At the current time t_o = 13:06:00 LST, the prior 10 minute GHI data is collected, and $t_o - 10$ min is defined as the initial time t_i = 12:56:00 LST. Since more than one large down ramp occurred in the ten minute window, the down ramps closest in time to t_i are selected and the ramp start time instants are determined. The detected down ramp start times are t = 12:59:46 LST for BMSB and t = 13:01:19 LST for EBU2 (red dots) yielding cloud travel times defined in Eq. (11**Error! Reference source not found.** of $\Delta t = |t - t_i|$ = 226 s for BMSB and 319 s for EBU2, respectively.

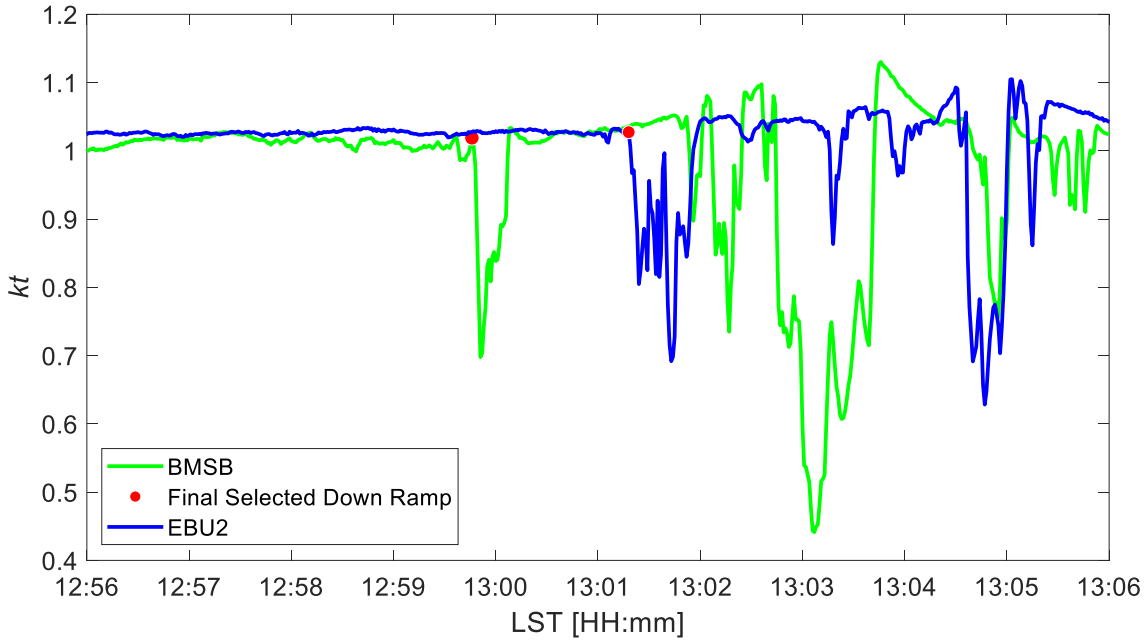


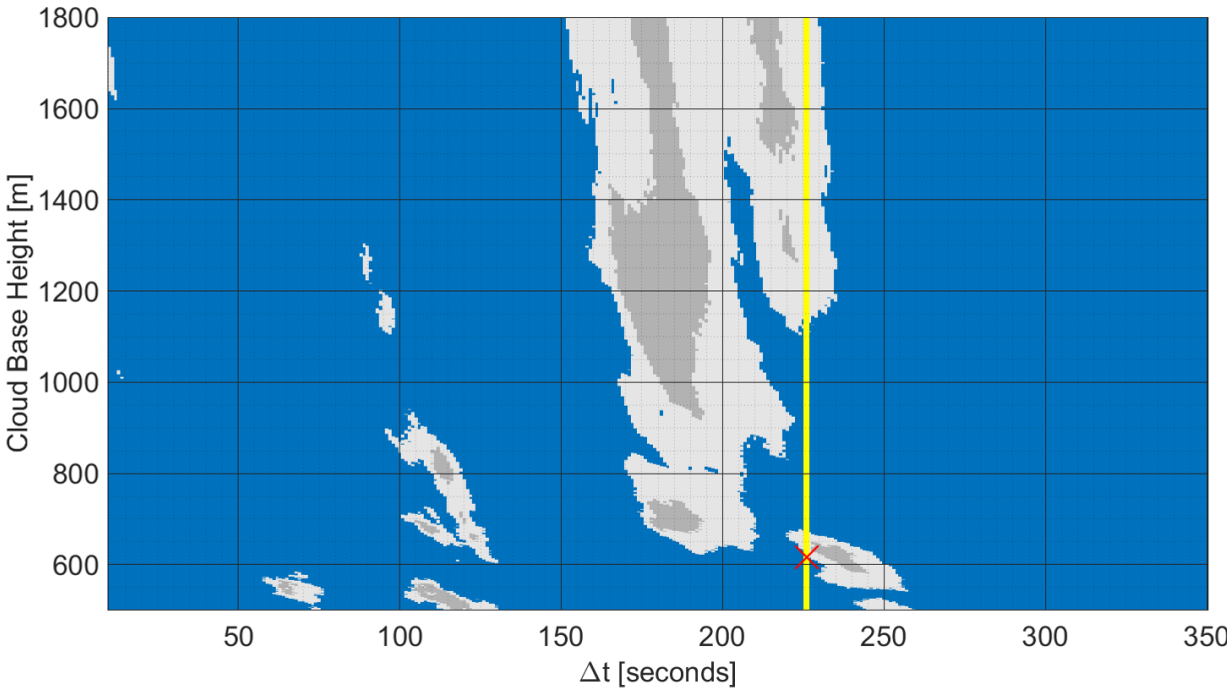
Figure 6: Illustration of the proposed ramp detection procedure to determine the start time of the down ramp for the BMSB and EBU2 station at initial time t_i = 12:56:00 LST on May 19, 2014. The start time of the final selected down ramp closest to t_i is marked as a red dot.

3.2.3 **Using the H-Δt Map to Estimate CBH**

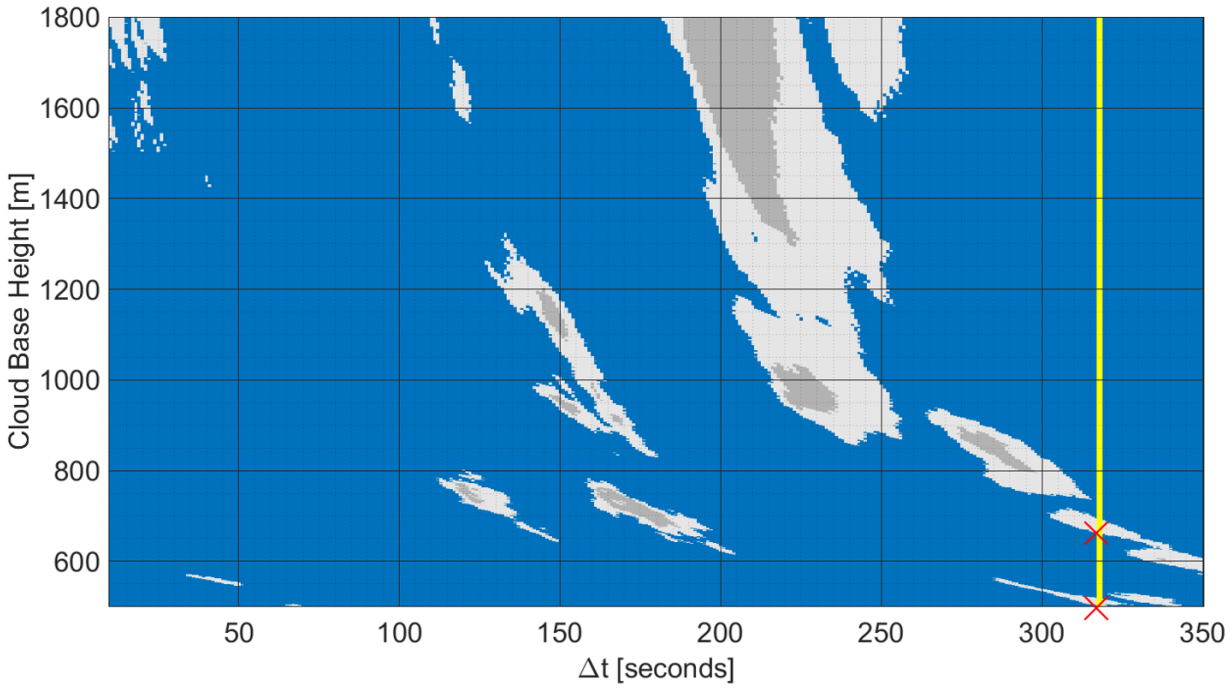
Equation (10 provides an expression for cloud map pixel location as a function of cloud base height H, cloud travel time Δt, and cloud velocity $\mathbf{u}$. The cloud state of the cloud map at location $\hat{\mathbf{x}}_c(H, \Delta t)$ can be clear sky, thin cloud, or thick cloud. The range of H considered in our analysis is 300 m to 2500 m in 50 m increments based on the common CBH range for coastal Southern California. CBHs are limited to 2,500 m as 5 years of CBH measurements from 12 METAR stations in southern California showed that 93% of CBHs are below 2500 m (not shown). Δt is varied from 0 to 10 min in 5 sec increments. Using the grid of H and Δt (velocity is assumed constant during Δt), the pixel position in the cloud map is computed (Eq. (10), and the cloud state is extracted. This results in a transformation of the cloud map which we call H-Δt map.

Figure 7 visualizes the H-Δt map for the time window and GHI data corresponding to Figure 6. For illustration purposes, the CBH range is set to 500 m to 1800 m in 10 m increments and Δt varies from 0 to 350 sec in 5 sec increments. The vertical yellow lines are placed at Δt_{BMSB} and Δt_{EBU2}, indicating the respective station cloud travel times (as determined in Fig. 6). CBH candidates are obtained from the H-Δt map by searching for cloud condition transitions around

lines of constant Δt_i. The most commonly occurring CBH candidate across all stations is selected as the CBH estimate. If two or more CBH candidates are equally common then they are averaged. If none of the stations returns a CBH candidate, no CBH estimate is generated. Red crosses in Fig. 7 indicate the CBH candidates are 620 m for BMSB, and 540 m and 660 m for EBU2. Thus, the CBH candidates from the two stations are averaged to be 606 m. The concurrent ceilometer reading at $12{:}\,59{:}\,00$ LST indicates a single cloud layer at 610 m.



(a)



(b)

Figure 7: CBH versus arrival time (or H- Δt map) for the initial time $t_i = 12{:}56{:}00$ LST on May 19, 2014. Ramp events, indicated by the vertical yellow lines, were detected at $\Delta t = 226$ s and $\Delta t = 319$ s for the (a) BMSB and (b) EBU2 ground stations, respectively (see Figure 2 for locations and Figure 6 for ramp detection). Left-to-right transitions from clear to cloudy (i.e. down ramps) along the yellow line indicate CBH candidates (red cross). Blue, white, and grey colors represent clear sky, thin clouds, and thick clouds, respectively.

Note that the Δt axis scales linearly with the cloud velocity, so uncertainty in $\mathbf{u}$ contributes directly to uncertainty in Δt. This and other potential errors in down ramp timing estimates (see discussion in Section 5.2) justify extending the cloud condition transition search to a search window of 60 seconds around lines of constant Δt_i. While this process induces more CBH candidates, it reduces the times when no CBH estimate is output.

An alternate, more intuitive presentation of the H- Δt map is the cloud shadow distribution plan view in Fig. 8, generated using the t_i cloud map advected at cloud velocity $\mathbf{u}$ to time $t_i + \Delta t$. In each subplot of Fig. 8, a shadow is just about to pass over the stations (note the shadow adjacent to each red dot and the direction of cloud movement). The plan view gives the cloud shadow distribution for all stations at a single $(H, \Delta t)$ pair, whereas the H- Δt map gives the possible cloud

condition for a single station at a range of $(H, \Delta t)$.

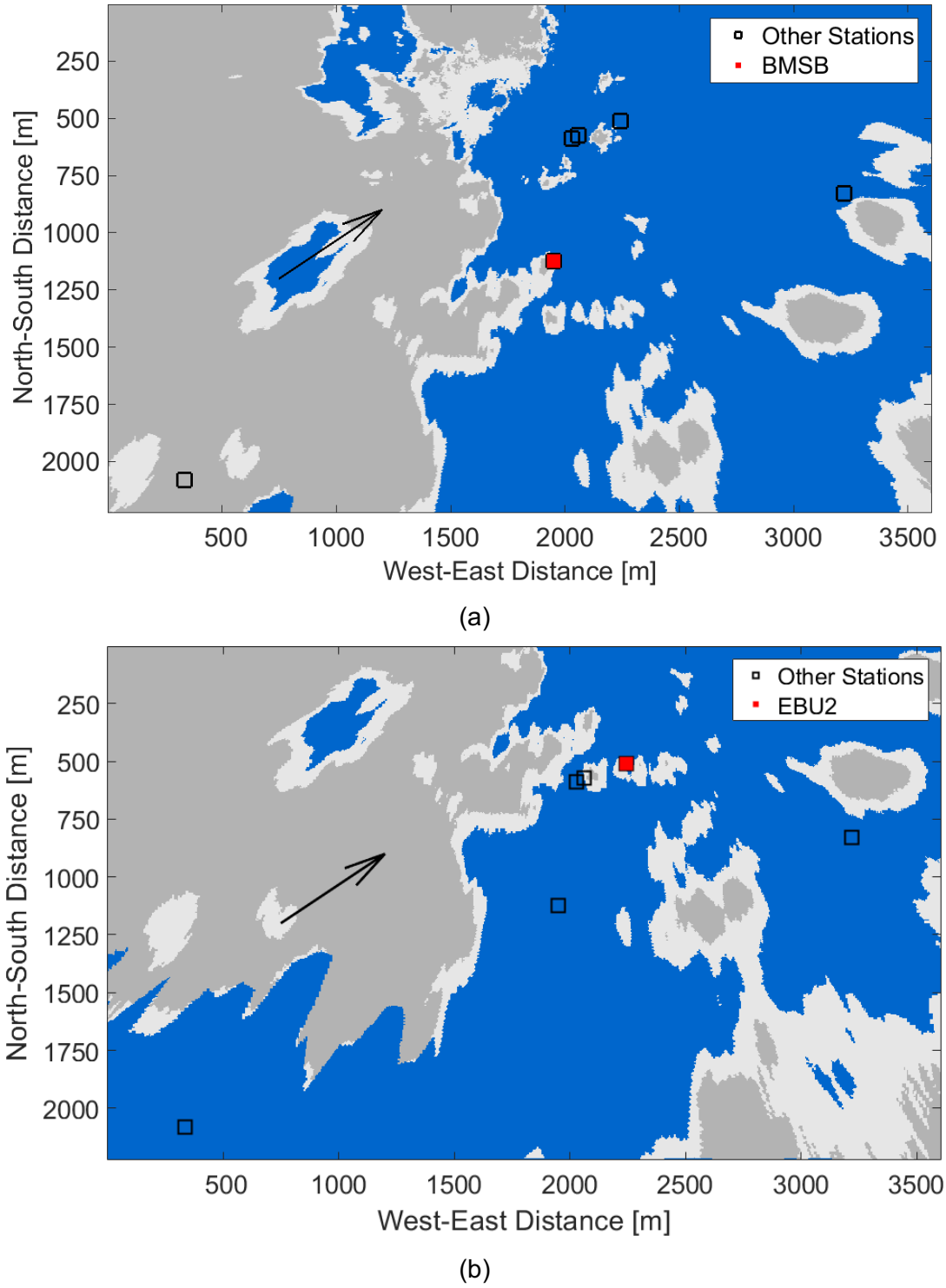


Figure 8: Advected cloud shadow map generated from the sky image taken at $t_i = 12{:}56{:}00$ LST on May 19, 2014 using arrival time of the down ramp of (a) $\Delta t_{\mathrm{BMSB}} = 226$ s and (b) $\Delta t_{\mathrm{EBU2}} = 319$ s and a

CBH of 606 m determined in Fig. 7. BMSB (a) and EBU2 (b) ground stations are shown as red filled squares. Empty squares represent the five other ground stations. The arrow indicates the cloud motion vector, showing the cloud shadows moving towards northeast. The arrow magnitude indicates the distance traveled by a cloud in 30 s. Blue, white, and grey colors represent clear sky, thin cloud, and thick cloud, respectively.

4. Results

4.1 **Median Filtering and Error Metrics**

The non-uniform ceilometer measurements are first resampled to the TSC and GCSE time steps through nearest neighbor interpolation. A sliding 15 minute median filter is then applied to the raw output of TSC, GCSE, and the resampled ceilometer measurements. To quantify the differences between the proposed methods and the ceilometer output, the mean bias difference (MBD) and the root mean square difference (RMSD) were used:

$$\mathrm{MBD} = \frac{1}{\boldsymbol{N}}\sum_{\boldsymbol{n}=1}^{\boldsymbol{N}}(H_{\boldsymbol{n}}^{\mathrm{model}} - H_{\boldsymbol{n}}^{\mathrm{ceil}}), \tag{12}$$

$$\mathrm{RMSD} = \sqrt{\frac{1}{\boldsymbol{N}}\sum_{\boldsymbol{n}=1}^{\boldsymbol{N}}(H_{\boldsymbol{n}}^{\mathrm{model}} - H_{\boldsymbol{n}}^{\mathrm{ceil}})^2}, \tag{13}$$

where $\boldsymbol{N}$ is the total number of data points, $H_{\boldsymbol{n}}^{\mathrm{model}}$ is the CBH from the TSC and GCSE methods, and $H_{\boldsymbol{n}}^{\mathrm{ceil}}$ is the corresponding ceilometer measurement at time index $\boldsymbol{n}$. MBD and RMSD are divided by the daily average CBH measurement from ceilometer to obtain a normalized MBD (nMBD) and normalized RMSD (nRMSD). Normalization provides a better comparison across days (RMSD is expected to be proportional to the true cloud height), whereas the un-normalized metrics give a better characterization of CBH accuracy for solar power forecasting. Periods with rain, either falling or droplets remaining on the sky imager, were excluded from the evaluation since neither ceilometer nor sky imager methods perform reliably under those conditions. Rainy periods are shaded in yellow.

4.2 **Evaluation over 30 Cloudy Days**

The performance for 30 days, spanning all seasons and multiple cloud types is given in Figure

9 and summarized in Table 2 (see Table A-1 in Appendix for complete comparison). Stratocumulus and cumulus clouds were most common on the selected days. Only four of the 30 days had CBHs exceeding 1000 m, so the evaluation provided is predominantly for low cloud conditions consistent with the dominant climatology of coastal Southern California. Overall, TSC outperformed GCSE for this extended data set, with TSC achieving an average RMSD of 133 m versus 163 m for GCSE. The standard deviation of daily RMSD for TSC was 72.3 m versus 92.9 m for GCSE, indicating the performance of TSC is more consistent across days. TSC had a small positive bias, versus a small negative bias for GCSE.

The number of CBH values reported per day varies markedly between TSC and GCSE. GCSE yields no result if there are no clouds detected that will shade the station. This will occur during periods with sufficiently homogenous cloud conditions and specifically periods with clear or overcast conditions along the cloud motion vector such that no CBH candidates are generated for the available ground stations, i.e. there are no clouds within $\Delta t_i \pm \sigma_t$ for each station i in the H-Δt map. Additionally, GCSE cannot generate CBH if no down ramps are located. These limitations cause GCSE to issue 34% less CBH than TSC averaged over 30 days.

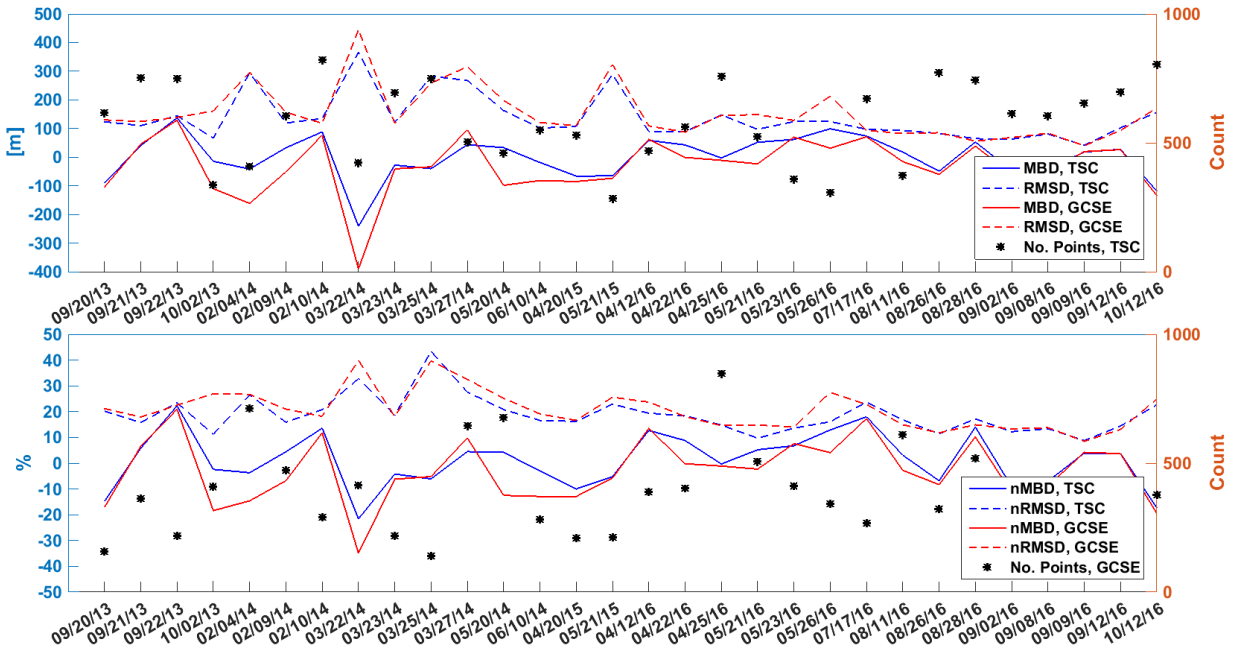


Figure 9: Validation of cloud base height estimates for 30 days. Line styles distinguish error metrics, and line colors differentiate TSC and GCSE methods, respectively. The number of raw measurements are displayed in black dots (right y axis).

Table 2: The monthly average (Avg.) and standard deviation (Std.) of the daily error metrics weighted

according to the number of data points. Ceilometer daily averages are reported as 'Mean CBH'. Refer to Table A-1 in the Appendix for error metrics by day.

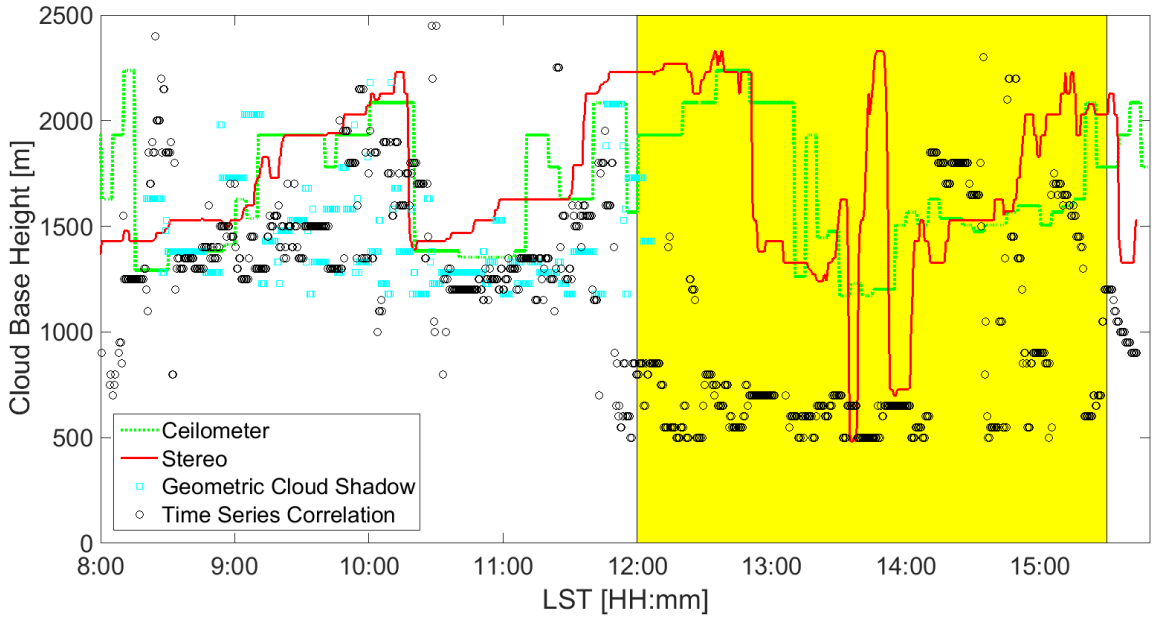
	Mean CBH [m]	TSC method					GCSE method				
		MBD [m]	nMBD [%]	RMSD [m]	nRMSD [%]	No. Points	MBD [m]	nMBD [%]	RMSD [m]	nRMSD [%]	No. Points
Avg.	715	2.2	1.3	132.8	18.9	581	-33.9	-3.1	162.6	20.8	382
Std.	217	72.0	-	72.3	-	160	101	-	92.9	-	176

4.3 **Comparison to NK14 on Select Days**

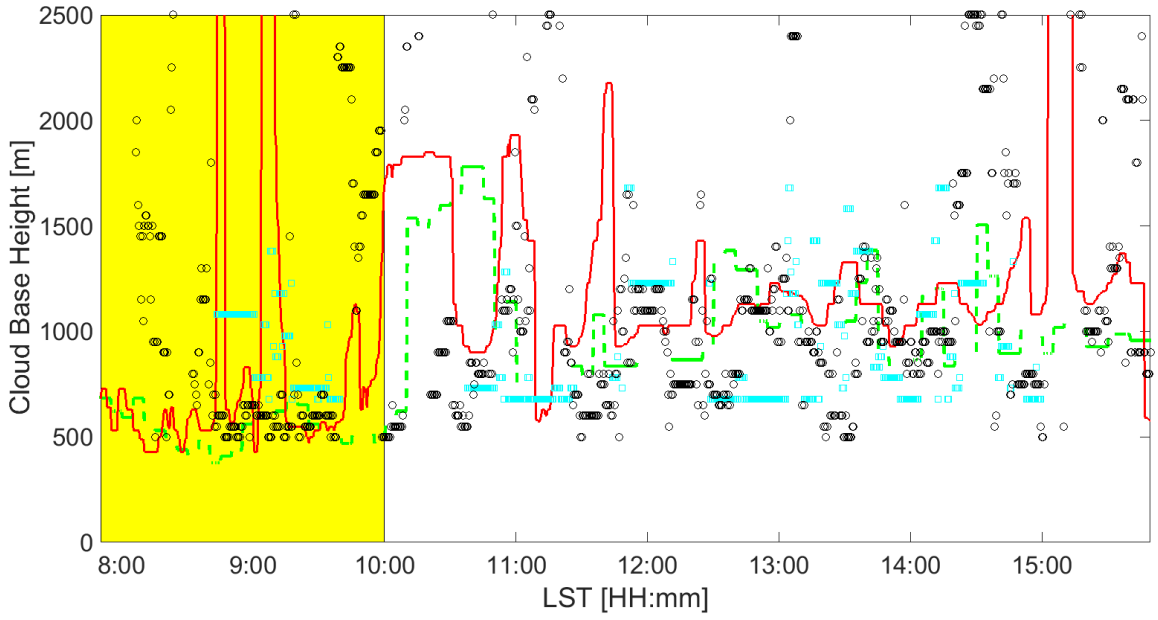
Table 3 and Figure 10 present further validation against NK14 on three days. While it produces scattered raw results, TSC captures the major CBH transition on all three days. In contrast, the CBH estimates from GCSE are not as scattered likely because of the internal quality control that requires CBH output consensus between stations. RMSD errors for TSC and NK14 are less than 300 m (RMSD) and 20% (nRMSD) averaged over the three days. GCSE, however, has RMSD and nRMSD of over 400 m and 27%, respectively, performing consistently worse than the other two methods. The MBD and nMBD show that the bias of GCSE is almost twice that of TSC for these three days. Note that nRMSD seems higher for Dec 26 on both methods; however, the absolute error on Dec 26 is not unusual and the large error can be attributed to the normalization by a smaller CBH. NK14 beats both TSC and GCSE on all three days, though the performance of TSC is close to that of NK14.

Table 3: Comparison of cloud base height estimates. Ceilometer daily averages are reported as 'Mean CBH'. The average for each column ('Avg.') is weighted according to the number of data points in each day. Rainy periods are excluded.

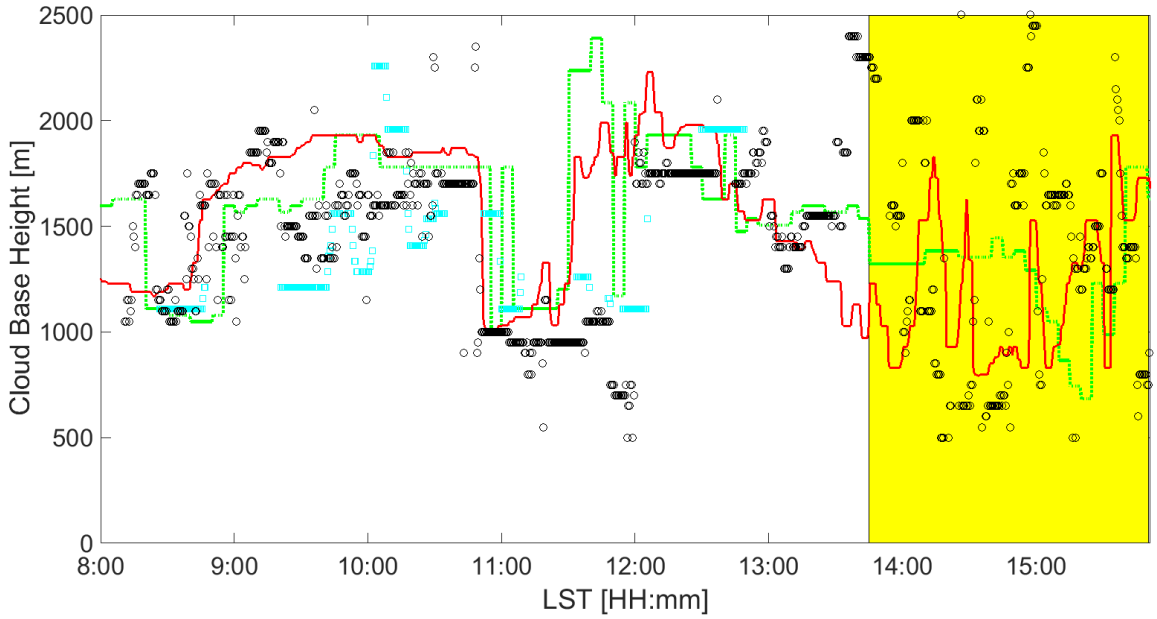
	Mean	TSC method				GCSE method				NK14: Stereographic	
Date	CBH [m]	MBD [m]	nMBD [%]	RMSD [m]	nRMSD [%]	MBD [m]	nMBD [%]	RMSD [m]	nRMSD [%]	RMSD [m]	nRMSD [%]
Dec 14	1814	88	4.9	291	16.0	-223	-12.3	394	21.7	262	14.3
Dec 26	1164	-128	-11.0	288	24.8	-323	-27.8	399	34.3	206	17.7
Dec 29	1625	-103	-6.3	299	18.4	169	10.4	440	27.1	272	16.8
Avg.	1534	-47.8	-4.0	293	19.5%	-104.3	-8.3	413	27.4	246	16.3


Cloud Base Height [m]
2500
2000
1500
1000
500
0
8:00
9:00
10:00
11:00
12:00
13:00
14:00
15:00
LST [HH:mm]
Ceilometer
Stereo
Geometric Cloud Shadow
Time Series Correlation

(a)


Cloud Base Height [m]
2500
2000
1500
1000
500
0
8:00
9:00
10:00
11:00
12:00
13:00
14:00
15:00
LST [HH:mm]

(b)



(c)

Figure 10: Cloud base height comparison between the TSC (black dot), GCSE (green dot), and 2D stereographic method (red, Nguyen and Kleissl, 2014), and ceilometer measurements (green dashed) for (a) Dec 14, (b) Dec 26, (c) Dec 29, 2012. Yellow highlights show periods of rain that are ignored in the error calculation in Section 4.1.

5. Discussion

5.1 **TSC Performance**

TSC computes the average correlation coefficient between 20 minutes of measured and modeled GHI across several ground stations. Correlation coefficients are computed for a range of CBH values, and the CBH corresponding to the maximum correlation is output. While the bias of the method over 30 days is small at 1.3% nMBD, the random error is significant at 18.9% nRMSD. Although this may seem high, it is within 3 percentage points of the stereographic method. The following subsections highlight different factors affecting the performance of TSC method.

5.1.1 **GHI Sampling and Correlation**

Forty-one (41) samples are used to compute the correlation coefficients which are subsequently averaged across stations. The 20 minute sample duration may be insufficient to

yield a reliable CBH estimate, but is chosen empirically to allow the method to be sufficiently dynamic to track intra-hour changes in CBH. Increasing the time window may reduce the estimator variance at the expense of being unable to react to rapid CBH changes. An alternative to increasing the sample duration is to decrease the sampling (image capture) time step (i.e. 30 seconds) in this work. Beyond increasing the number of samples, treating the sensor network as an array and applying array signal processing methods may provide a lower variance CBH estimator.

5.1.2 **Sensitivity of CBH to Correlation Coefficient**

During certain periods, the variation of the correlation coefficients R_j over all CBH candidates was found to be small. For example, Figure 11 gives the mean correlation R_j (Eq. **Error! Reference source not found.**) at different H_j for a selected period. The maximum R_j is very similar to the minimum with R_j ranging from 0.9 to 1. While the changes in R_j are small relative to its range, the relative changes in H_j are considerable at 1050 m to 1700 m. In this case, due to the small difference between the minimum and maximum correlation coefficient, the selected H_j may be determined by small and somewhat random fluctuations in R_j which is not desirable behavior for an accurate and robust CBH estimation algorithm. Small variations in the correlations are caused by homogenous cloud cover (e.g. overcast condition) or a cloud projection that is insensitive to CBH changes (e.g. collocated sky imagery and pyranometer).

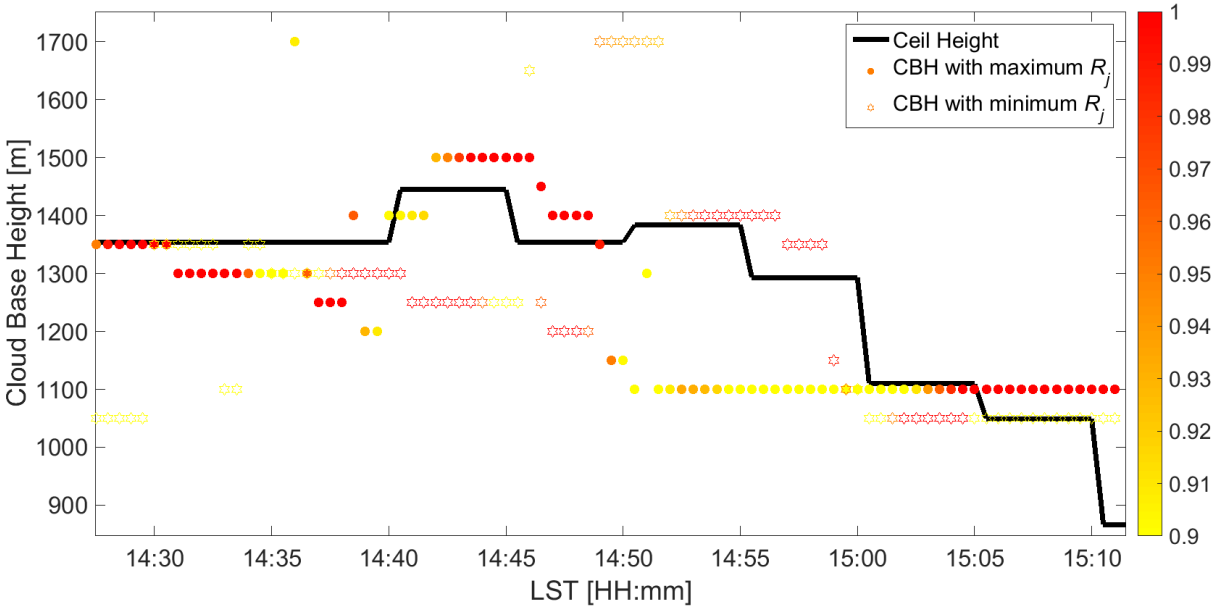


Figure 11: Example of CBH estimates for the TSC method versus the ceilometer for a 45 minute period on Dec 29, 2012. The color of each symbol indicates the average correlation coefficient R_j (Eqs. **Error! Reference source not found.** and **Error! Reference source not found.**) between the observed and simulated nowcast GHI from the set of stations. For each time step the CBH (*y*-axis) and its associated maximum R_j (filled circles) and minimum R_j (open hexagrams) are shown.

Moreover, larger imager to station distance can promote errors in the GHI time series from the sky imager. As indicated in Eq. (9 a larger pixel zenith angle (more distant station) results in cloud projection being more sensitive to CBH changes because the cloud projection error scales with $tan\theta$. In addition, the lower pixel resolution for the outer part of the sky image at larger pixel zenith angle can cause larger random errors in shadow projection at the ground station.

501

502 ## 5.2 **GCSE Performance**

503 GCSE combines ramp detection with an analytic-geometric component derived from the sky
504 imager forecast. Down ramp events are detected for each ground station and associated cloud
505 edges are matched in each station's H-Δt map. Since the construction of H-Δt map is a matrix
506 indexing operation for an image, it takes less than a second to construct H-Δt map on a typical
507 i5-powered workstation, making operational use feasible.
508 In terms of nRMSD, GCSE performed over 10 percentage points worse than the NK14 method
509 over the three days studied. In the more extensive 30 day comparison, GCSE improved
510 substantially with an nRMSD of 20.8%. For all error metrics, the GCSE performed worse than
511 TSC. This is in large part due to the modeling complexity and assumptions involved (see Section
512 5.3).
513

514 ### 5.2.1 **Down Ramp Start Time**

515 An accurate down ramp start time from GHI observations is required for the GCSE to work
516 correctly. The method described in Section 3.2.2 is a reasonable approach if the ramp is
517 monotonically up or down. But in some cases ramps exhibit local extrema, causing the proposed
518 approach to misidentify the start time.
519 Figure 12 provides an example on Dec 26 with scattered cumulus clouds. Fig. 12a shows a
520 large down ramp with a complex *kt* time series: two local extrema are identified in the time series

maximum ramp points difference (Figure 5) at 12:23:21 LST (black dashed line) and 12:23:52 LST (green dashed line). The associated ramp event start times are determined at 12:23:26 LST (black dot) and 12:23:47 LST (green dot), respectively, by searching within 5 second for a local maximum in *kt*. While visual inspection suggests that the black dot is a reasonable ramp start time, the *kt* variation around the two original local extrema is small, making identifying the start time somewhat random. These small "pre-ramp" events are likely caused by the multiscale nature of clouds and associated deformations around the cloud boundary.

The impact of this ambiguity in the local extremum is illustrated in Fig. 12b. The black dot in (a) corresponds to black line $(H, \Delta t)$ = (775 m, 86 s) and the green dot in (a) corresponds to green line $(H, \Delta t)$ = (625 m, 107 s) with Δt = 0 at 12:22:00 LST. The two local extrema that are spaced by 31 s cause a 150 m difference in CBH. In this case, the local extremum #1 is slightly greater than #2, so it is selected per the procedure in Section 3.2.2 and the associated local maximum (i.e. black dot) is used to determine the CBH candidate in Figure 12b. While in this case the final selected CBH is closer to the ceilometer measurement of 866 m than the alternate, similar ambiguities in local extremum and subsequent CBH variation were common in the analysis.

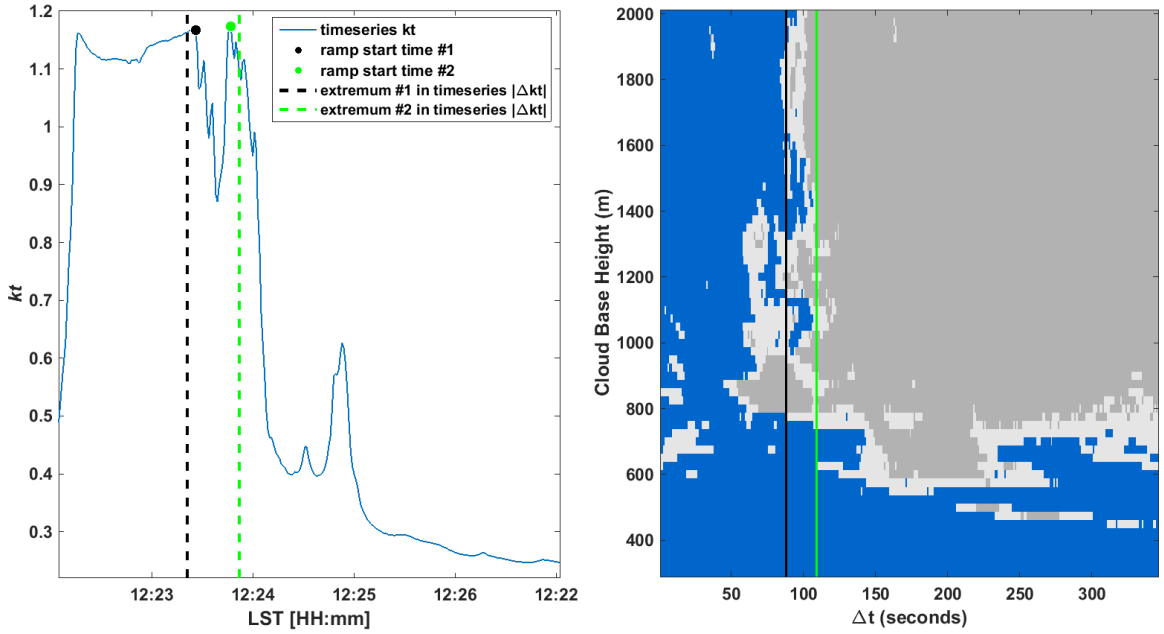


Figure 12: Sensitivity of CBH to ramp start time and ambiguity in ramp start time estimation. (a) Two local extrema (dashed lines) are identified due to a non-monotic time series of *kt* (the dots show how following ramp detection each ramp start time is adjusted to the local maximum in *kt* within 5 sec). (b) H-Δt map corresponding to (a) with Δt = 0 sec at 12:22:00 LST. The vertical lines in (b) correspond to

the colored dots in (a). The actual CBH measurement from ceilometer is 866 m.

5.3 **Other Modeling Errors Affecting CBH Estimation**

Both TSC and GCSE rely on derived products generated in the USI forecast procedure that apply simplifying assumptions and inject additional uncertainty into CBH estimation. Naturally, since sky images are the key input to both methods, TSC and GCSE are not operational at night. Cloud edges derived from sky imagery rely on the cloud decision process determining where clouds "begin". The methods to detect cloud presence are generally accurate, but there is some inherent uncertainty in a binary pixel classification as being-cloudy or cloud-free (Ghonima et al., 2012), particularly near cloud edges which may have a diffuse and blurred transition. This affects both TSC and GCSE.

Another issue is that extensive cloud evaporation and formation can cause GCSE to fail because the “frozen” cloud advection assumption is violated. Consider the case where a cloud forms between time t_o when a sky image is taken, and the time when that cloud’s edge causes a down ramp at time t_1. Another stable cloud that was present in the sky image at t_o causes a down ramp at time t_2 where $t_2 > t_1$. Although the down ramp occurring at t_1 is detectable in GHI data, the cloud map generated from data at t_o only has the information of the cloud which passes at t_2. For TSC, this increases the separation between the measured and modeled GHI time series, affecting correlation coefficients across the CBH grid search. The GCSE ramp detection algorithm will return a ramp occurrence time of $\Delta t_1 = t_1 - t_o$ which does not have a matching cloud edge in the $H\text{-}\Delta t$ map. The $H\text{-}\Delta t$ map search process will yield the best available clear-cloudy transition at Δt_1 which is likely to be incorrect.

Besides, both methods are affected by overcast conditions with homogenous cloud cover. The TSC method identifies concurrent cloud edge events using the correlation coefficient. If the 20 minute sample window does not contain any significant cloud-edge induced fluctuations, the correlation coefficients are small and likely no CBH will be output. The GCSE method does not provide CBH in overcast conditions either; while ramp detection may still be feasible due to variability of cloud optical depth in overcast conditions, the cloud travel time cannot be estimated from the binary $H\text{-}\Delta t$ map. Fortunately, in overcast or clear conditions the solar irradiance can be predicted accurately without CBH because all stations are likely covered by the same sky condition and receive similar irradiance.

Additionally, the pixel resolution for the outer part of the sky image at larger pixel zenith angle

is degraded, making the estimated cloud cover more uniform over the 20-minute comparison interval. Any station whose shadow projection comes from these perimeter image sections will lack detailed cloud structure. This less detailed cloud structure yields lower correlation for the TSC, and larger errors in identifying the timing of sky condition changes for the GCSE. Interestingly, for the GCSE, the forecast at the sky imager position is unaffected by CBH and thus the forecast GHI does not suffer from CBH errors.

The temporal resolutions for TSC and GCSE differ: the TSC output rate is one sample per 30 seconds as set by the image capture frequency, but data availability may be less frequent due to low correlation coefficients. The GCSE's output rate depends on the existence of sufficient variability in cloud cover, and the ability to find a consensus CBH candidate. For the dataset presented here, GCSE outputs 34% less CBH samples than TSC, for an average of one GCSE sample every 75 seconds. While this lower output rate is sufficient for short-term solar power forecasting, it may be a limiting factor for other applications or in other sky conditions with less cloud cover. The valid time of the CBH estimates also differs between methods: Since TSC correlates the last 20 minutes of GHI data, the estimated CBH applies to those 20 minutes. While GCSE utilizes only a single dominant down ramp in the GHI time series the CBH strictly applies to that time instant only.

The cloud velocity estimation of the sky imager is actually an apparent cloud edge velocity, which is a combination of cloud speeds due to advection along with cloud formation or evaporation occurring from image to image. These cloud dynamics introduce real or apparent fluctuations in cloud speed which negatively affects the performance of GCSE because construction of the H-Δt map assumes that the cloud velocity remains constant over the CBH estimation interval (typically 10 minutes). TSC is insensitive to cloud speed variability as it does not employ a cloud advection scheme.

Last, multiple cloud layers and cloud three-dimensionality (Mejia et al., 2018) can degrade the performance because both methods operate under assumption of single-layered planar cloud cover.

5.4 **Number of Stations and Spatial Diversity**

Geographic variations at the individual sites may affect both TSC and GCSE. For TSC, averaging correlation coefficients at each CBH blurs potential station-to-station differences in correlation coefficient due to real differences in CBH. For GCSE, station-to-station ramp timing errors may cause inconsistent CBH candidates, preventing an accurate CBH estimate. The current limitation of our setup was the availability of only six stations, four of which were located

within 600 m of each other resulting in more correlated GHI data and little diversity in perspectives. A logical extension to this work is to examine the impact of adding additional ground stations. At large solar installations, weather stations, reference cells, and individually metered inverters can all be used to improve spatial distribution of stations.

6. Conclusions and Future Work

The objective of this paper was to propose two methods for CBH estimation requiring a single sky imager together with spatially distributed irradiance or power output measurements, providing an alternative CBH estimation technique to direct, in-situ, or multi-camera approaches. These new methods can serve as a low-cost alternative to ceilometers for sky imager based short-term solar power forecasting in which the cloud height information is required (Chow et al., 2011; Schmidt et al., 2016).

The TSC method, is comparatively simple and the more reliable of the two proposed methods. The GCSE method relies on a complex stack of models: cloud detection, cloud velocity estimation, cloud shadow forecasting, and down ramp detection. The construction of the H-Δt map is a novel feature of this work, and its utility is demonstrated for the purposes of cloud edge matching and CBH estimation. Overall, the GCSE method performed slightly worse (1 percentage point larger nRMSD) than the TSC method. For both methods, the nRMSD remained below 21% for all 30 days. On the other hand, the CBH estimate derived from a sky imager coupled with a cloud speed sensor in our previous work (Wang et al., 2016) yielded better accuracy (17% nRMSD) on a different set of 30 days, owing partially to the strict filtering of the raw cloud speed measurement.

Future efforts will involve improving both sky imager cloud detection and cloud velocity estimation, which will also benefit solar power forecasting with a sky imager. Chow et al. (2015) proposed optical flow to enable detection of multiple cloud layers as well as their respective cloud pixel speeds, which is an improvement to the cross-correlation velocity estimation method used in this work. Adding more and more distributed ground stations will also help improve the robustness of the methods. Finally, validation under different meteorological conditions more relevant to continental climates would further substantiate the general applicability of the methods.

Appendix

Table A-1: Comparison of cloud base height estimates between TSC and GCSE methods for 30 days. Ceilometer daily averages are reported as 'Mean CBH'.

Date	Mean CBH [m]	TSC method					GCSE method				
		MBD [m]	nMBD [%]	RMSD [m]	nRMSD [%]	No. Points	MBD [m]	nMBD [%]	RMSD [m]	nRMSD [%]	No. Points
09/20/13	613	-90.5	-14.8	123.6	20.2	616	-104.7	-17.1	129.9	21.2	157
09/21/13	695	40.6	5.8	109.9	15.8	752	45.3	6.5	124.4	17.9	362
09/22/13	618	137.0	22.2	144.5	23.4	750	129.0	20.9	138.9	22.5	217
10/02/13	598	-14.5	-2.4	66.7	11.2	337	-109.9	-18.4	161.1	26.9	408
02/04/14	1110	-41.5	-3.7	294.4	26.5	410	-161.6	-14.6	296.0	26.7	712
02/09/14	747	32.7	4.4	118.9	15.9	606	-51.1	-6.8	156.8	21.0	473
02/10/14	648	88.3	13.6	135.0	20.8	822	76.5	11.8	118.2	18.2	291
03/22/14	1116	-241.4	-21.6	365.9	32.8	424	-389.6	-34.9	444.7	39.9	416
03/23/14	657	-27.7	-4.2	122.3	18.6	696	-40.8	-6.2	118.6	18.1	219
03/25/14	651	-40.0	-6.1	282.6	43.4	750	-33.7	-5.2	258.7	39.7	141
03/27/14	971	43.5	4.5	268.0	27.6	504	94.9	9.8	315.7	32.5	646
05/20/14	783	33.1	4.2	162.7	20.8	460	-98.3	-12.5	197.7	25.2	679
06/10/14	633	-18.6	-3.0	101.7	16.6	552	-81.6	-12.9	120.8	19.1	283
04/20/15	662	-65.9	-10.0	107.2	16.2	531	-85.1	-12.9	110.2	16.6	209
05/21/15	1263	-65.6	-5.2	288.8	22.9	285	-73.8	-5.8	324.0	25.6	213
04/12/16	460	58.6	12.7	89.4	19.4	471	62.1	13.5	108.9	23.7	389
04/22/16	485	42.7	8.8	89.4	18.4	563	-0.9	-0.2	87.6	18.1	404
04/25/16	998	-2.8	-0.3	147.4	14.8	760	-11.1	-1.1	145.7	14.6	847
05/21/16	1004	51.9	5.2	97.1	9.7	525	-23.5	-2.3	148.4	14.8	507
05/23/16	916	61.9	6.8	124.4	13.6	361	69.7	7.6	128.4	14.0	411
05/26/16	771	99.2	12.9	124.3	16.1	309	31.3	4.1	212.4	27.5	342
07/17/16	409	73.5	18.0	97.1	23.7	672	70.9	17.3	93.3	22.8	266
08/11/16	555	17.4	3.1	93.4	16.8	374	-15.7	-2.8	82.9	14.9	610
08/26/16	725	-49.0	-6.8	82.5	11.4	773	-60.0	-8.3	85.3	11.8	323
08/28/16	376	52.8	14.0	64.5	17.2	744	38.4	10.2	56.0	14.9	520
09/02/16	514	-46.6	-9.1	62.9	12.2	614	-59.3	-11.5	68.4	13.3	280
09/08/16	603	-42.5	-7.0	80.6	13.4	605	-53.5	-8.9	83.4	13.8	253
09/09/16	469	18.1	3.9	41.2	8.8	654	19.5	4.2	40.0	8.5	278
09/12/16	714	27.3	3.8	102.7	14.4	699	26.3	3.7	93.2	13.1	237
10/12/16	689	-118.9	-17.3	156.4	22.7	805	-134.4	-19.5	171.5	24.9	377

References

Allmen, M. C., & Kegelmeyer Jr, W. P. (1996). The computation of cloud-base height from paired whole-sky imaging cameras. *Journal of Atmospheric and Oceanic Technology, 13(1), 97-113*.

Bosch, J.L & Kleissl, J. (2013). Cloud motion vectors from a network of ground sensors in a solar power plant. *Solar Energy, volume 95, pp. 13-20*.

Bosch, J.L, Zheng, Y., & Kleissl, J. (2013). Deriving cloud velocity from an array of solar radiation measurements. *Solar Energy, volume 87, pp. 196-203*.

Bright, J., Smith, C., Taylor, P., & Crook, R. (2015). Stochastic generation of synthetic minutely irradiance time series derived from mean hourly weather observation data. *Solar Energy, 115, 229-242*.

Cazorla Cabrera, A. (2010). Development of a sky imager for cloud classification and aerosol characterization. Universidad de Granada. Retrieved from http://hdl.handle.net/10481/5533

Chow, C. W., Belongie, S., & Kleissl, J. (2015). Cloud motion and stability estimation for intra-hour solar forecasting. *Solar Energy, 115, 645-655*.

Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., et al. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy, 85(11), 2881-2893*.

Dessler, A., Palm, S., & Spinhirne, J. (2006). Tropical cloud top height distributions revealed by the ice, cloud, and land elevation satellite (ICESat)/Geoscience laser altimeter system (GLAS). *Journal of Geophysical Research: Atmospheres (1984–2012), 111(D12)*.

Fung, V., Bosch, J., Roberts, S., & Kleissl, J. (2014). Cloud shadow speed sensor. *Atmospheric Measurement Techniques, 7(1), 1693-1700.*

Gaumet, J., Heinrich, J., Cluzeau, M., Pierrard, P., & Prieur, J. (1998). Cloud-base height measurements with a single-pulse erbium-glass laser ceilometer. *Journal of Atmospheric and Oceanic Technology, 15(1), 37-45.*

Ghonima, M., Urquhart, B., Chow, C., Shields, J., Cazorla, A., & Kleissl, J. (2012). A method for cloud detection and opacity classification based on ground based sky imagery. *Atmospheric Measurement Techniques, 5(11), 2881-2892.*

Hutchison, K., Wong, E., & Ou, S. C. (2006). Cloud base heights retrieved during night-time conditions with MODIS data. *International Journal of Remote Sensing, 27(14), 2847-2862.*

Ineichen, P., & Perez, R. (2002). A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy, 73(3), 151-157. doi: 10.1016/S0038-092X(02)00045-2*

Kassianov, E., Long, C. N., & Christy, J. (2005). Cloud-base-height estimation from paired ground-based hemispherical observations. *Journal of Applied Meteorology, 44(8), 1221-1233.*

Killius, N., Prahl, C., Hanrieder, N., Wilbert, S., and Schroedter-Homscheidt, M. (2015) On the use of NWP for Cloud Base Height Estimation in Cloud Camera-Based Solar Irradiance Nowcasting. *ICEM 2015, 23-26. June 2015, Boulder, USA.*

Kleissl, J., Urquhart, B., Ghonima, M., Dahlin, E., Nguyen, A., Kurtz, B., Chow C.W., and Mejia, F.A. (2016). University of California, San Diego (UCSD) Sky Imager Cloud Position Study Field Campaign Report. *United States Department of Energy, Atmospheric Radiation Measurement Program Report DOE/SC-ARM-15-056.*

672 Kuhn, P., Wirtz, M., Wilbert, S., Bosch, J. L., Wang, G., Ramirez, L., ... & Pitz-Paal, R. (2018a).
673 Field validation and benchmarking of a cloud shadow speed sensor. Solar Energy, 173, 229-
674 245.

675 Kuhn, P., Wirtz, M., Killius, N., Wilbert, S., Bosch, J. L., Hanrieder, N., ... & Heinemann, D.
676 (2018b). Benchmarking three low-cost, low-maintenance cloud height measurement systems
677 and ECMWF cloud heights against a ceilometer. *Solar Energy*.

678 Liu, L., Sun, X., Liu, X., Gao, T., & Zhao, S. (2015). Comparison of cloud base height derived from
679 a ground-based infrared cloud measurement and two ceilometers. *Advances in Meteorology,*
680 *2015.*

681 Martinez-Anido, C. B., Botor, B., Florita, A. R., Draxl, C., Lu, S., Hamann, H. F., & Hodge, B. M.
682 (2016). The value of day-ahead solar power forecasting improvement. *Solar Energy*, *129*,
683 192-203.

684 Martucci, G., Milroy, C., & O'Dowd, C. D. (2010). Detection of cloud-base height using Jenoptik
685 CHTSC5K and Vaisala CL31 ceilometers. *Journal of Atmospheric and Oceanic Technology,*
686 *27(2), 305-318.*

687 Marquez, R., & Coimbra, C. F. (2013). Intra-hour DNI forecasting based on cloud tracking image
688 analysis. *Solar Energy, 91, 327-336.*

689 Mejia, F. A., Kurtz, B., Levis, A., de la Parra, Í., & Kleissl, J. (2018). Cloud tomography applied to
690 sky images: A virtual testbed. *Solar Energy, 176, 287-300*.

691 Nguyen, D. A., & Kleissl, J. (2014). Stereographic methods for cloud base height determination
692 using two sky imagers. *Solar Energy, 107, 495-509.*

693 Prata, A., & Turner, P. (1997). Cloud-top height determination using ATSR data. *Remote Sensing*
694 *of Environment, 59(1), 1-13.*

695 Schmidt, T., Kalisch, J., Lorenz, E., & Heinemann, D. (2016). Evaluating the spatio-temporal
696 performance of sky-imager-based solar irradiance analysis and forecasts. *Atmospheric*
697 *Chemistry and Physics, 16(5), 3399-3412.*

698 Shaw, J. A., & Nugent, P. W. (2013). Physics principles in radiometric infrared imaging of clouds
699 in the atmosphere. *European Journal of Physics, 34(6), S111 – S121.*

700 Urquhart, B., Ghonima, M., Nguyen, D., Kurtz, B., Chow, C. W., and Kleissl, J. Solar Energy
701 Forecasting and Resource Assessment. *Elsevier, 2013. ISBN: 978-0-12-397177-7, doi:*
702 *10.1016/B978-0-12-397177-7.00009-7*

703 Urquhart, B., Kurtz, B., Dahlin, E., Ghonima, M., Shields, J., & Kleissl, J. (2015). Development of
704 a sky imaging system for short-term solar power forecasting. *Atmospheric Measurement*
705 *Techniques Discussions, 7, 4859-4907.*

706 Urquhart, B., Kurtz, B., & Kleissl, J. (2016). Sky camera geometric calibration using solar
707 observations. *Atmospheric Measurement Techniques, 9(9), 4279-4294.*

708 Wang, J., & Rossow, W. B. (1995). Determination of cloud vertical structure from upper-air
709 observations. *Journal of Applied Meteorology, 34(10), 2243-2258.*

710 Wang, G., Kurtz, B., & Kleissl, J. (2016). Cloud base height from sky imager and cloud speed
711 sensor. *Solar Energy, 131, 208-221.*

712 Yang, H., Kurtz, B., Nguyen, D., Urquhart, B., Chow, C. W., Ghonima, M., et al. (2014). Solar
713 irradiance forecasting using a ground-based sky imager developed at UC San Diego. *Solar*
714 *Energy, 103, 502-524.*

715